

ОТЗЫВ

официального оппонента на диссертацию Думина Павла Николаевича «Математическое моделирование и идентификация параметров адаптивного тестирования с учетом временной динамики выполнения заданий», представленную на соискание учёной степени кандидата физико-математических наук по специальностям 05.13.18 – Математическое моделирование, численные методы и комплексы программ, 05.13.01 – Системный анализ, управление и обработка информации (авиационная и ракетно-космическая техника)

1. Актуальность темы

В настоящее время тестирование представляет собой достаточно распространённый инструмент организации педагогических измерений. Особое место среди различных вариантов тестирования занимает адаптивное, предполагающее формирование списка предъявляемых вопросов/заданий в зависимости от успешности выполнения испытуемым предыдущих этапов тестирования. Широкое распространение данный вид тестирования приобрёл в связи с развитием компетентного подхода в образовании, при котором недостаточно использовать для измерения традиционные вопросы, а стоит задача более комплексной оценки компетенций. Подобная задача подразумевает не только оценку фактических знаний, но и оценку некоторых практических навыков, в частности по работе со сложной специальной техникой, транспортными средствами, оборудованием, где квалификация специалиста оценивается не только с точки зрения теоретических знаний, но и способностями работать как в штатном режиме, так и при возникновении аварийных ситуаций. При этом оценивается не только квалификация, но и психологическая подготовка специалиста.

При подготовке и тестировании специалистов, в целях развития определённых профессиональных компетенций используются не только системы тестирования, но и специальные тренажёры-симуляторы, имитирующие процессы управления и обслуживания специальной техники, транспортных средств и оборудования. Фактически, работа с такими симуляторами представляет собой геймификацию (англ. Gamification) процесса обучения и подготовки специалистов. Одновременно с тем, симуляторы используются и для сдачи квалификационных экзаменов, в ходе которых проводится оценка профессиональных и психологических качеств специалистов, и соответственно определение уровня их квалификации, т.е. по сути, идёт их ранжирование по уровню пригодности к работе. Такой подход в частности применяется для диагностики лётного состава военной и гражданской авиации.

Поскольку при оценке квалификации и психологической подготовки об испытуемом представляется разнородная информация, отражающая соответствующие личностные и профессиональные качества, то в таких ситуациях идёт речь о принятии решений о принадлежности испытуемого к определённым группам пригодности (профессиональной и психологической). В таких ситуациях в ходе оценки испытуемых применяются системы поддержки принятия решений (СППР). На практике значительная часть работ по оценке результатов тестирований, подведения итогов производится вручную, лицами принимающими решение. В частности такая ситуация распространена при сдаче квалификационных экзаменов на симуляторах.

Исходя из рассмотренных особенностей, диссертационная работа, посвящённая

проблеме идентификации параметров адаптивного тестирования в условиях выполнения тестовых заданий во времени, а также в виде работы на симуляторах представляется актуальной для достаточно широкого круга областей. При этом работа находится на стыке двух научных специальностей 05.13.01 – «Системный анализ, управление и обработка информации (авиационная и ракетно-космическая техника)» в части систем поддержки принятия решений, оптимизации набора тестовых вопросов, их идентификации; 05.13.18 – «Математическое моделирование, численные методы и комплексы программ» в части разработки систем тестирования на базе симуляторов, моделирования процессов прохождения тестирования.

2. Анализ содержания диссертации

Диссертационная работа включает в себя введение, четыре главы и заключение.

Во введении даётся краткий обзор проблем организации адаптивного тестирования при анализе когнитивных способностей. Проблемы связаны, прежде всего, с порядком представления тестовых вопросов, оценкой объективности теста, детекции и устранения артефактов тестирования и др. Особое внимание уделяется проблематике на основе принципов игровой диагностики, практикуемых в работе с симуляторами. На основе рассмотренных особенностей ставится задача разработки математических моделей и методов адаптивного компьютерного тестирования навыков и способностей с учётом динамики изменения оцениваемых показателей и трудности заданий во времени. Разрабатываемые модели лежат в основе инструментариев систем поддержки принятия решений и систем моделирования (симуляторов) применяемых в игровой диагностике.

В обзорной части введения рассматриваются различные модели тестирования, а именно адаптивное тестирование с использованием байесовских сетей, с использованием цепей Маркова, на основе нейросетевых моделей.

В первой главе рассмотрена модель тестирования, разработанная на основе цепей Маркова, где выполнение заданий рассматривается в виде состояний, между которыми осуществляются переходы с определёнными вероятностями, а результат выполнения теста как исход Марковского процесса. На основе представленной модели разработана методика тестирования оператора сложных технических систем, применимая как для тестирования за компьютером, так и при работе с симуляторами и специализированными стендами.

Во второй главе рассмотрены методы оптимизации последовательности предъявления тестовых вопросов, направленные на сокращение числа вопросов, а также представление альтернатив с учётом сложности вопросов и предотвращение возможного угадывания вариантов ответа. Практическая реализация процесса тестирования осуществляется с помощью специально разработанного программного комплекса, включающего модули тестирования и идентификации тестовых заданий, а также хранилища тестовых заданий, их характеристик, выборочных данных и параметров моделей тестирования.

В третьей главе приведены вероятностные модели оценки пригодности тестовых заданий для измерения определённых качеств испытуемых, оценки времени выполнения заданий для различных категорий испытуемых. Вероятностные модели рассмотрены как для традиционных тестирований, так и для тестирований, проводимых в игровой форме (на симуляторе). В качестве практического применения рассматривается моделирование выполнения теста Равена группой испытуемых,

проводится сравнение эмпирических и теоретических данных.

В четвёртой главе производится сравнение разработанного метода идентификации моделей тестирования в сравнении с существующими. Применяется метод полной дискретизации, градиентный метод, метод дискретизации значимых параметров. Для оценки эффективности рассмотренных методов производится численный эксперимент, с пулом тестовых заданий и Марковской моделью из 4 состояний, определяемых 6 значимыми параметрами. Численно получены зависимости количества генерируемых моделей тестирования от количества состояний и количества генерируемых моделей тестирования от количества значимых параметров.

В заключении приведены основные результаты, выносимые на защиту.

Приведённые в разделах диссертации результаты полностью соответствуют научным положениям, выносимым на защиту.

3. Практическая значимость

Практическая значимость работы заключается диссертационной работы заключается в разработке математического обеспечения принципиально новых инструментариев для оценки компетенций. Инструментарии предполагают настройку используемых математических моделей на основе результатов испытуемых с учётом временной динамики выполнения тестовых заданий. Областью применения разработанного математического обеспечения являются не только традиционные системы тестирования, но и симуляторы, предполагающие игровую диагностику.

4. Достоверность

Достоверность полученных результатов не вызывает сомнений и подтверждается оценкой адекватности полученных результатов наблюдениям с помощью статистических критериев согласия. Практическим применением разработанного математического обеспечения в качестве инструментария систем поддержки принятия решений. Вычислительные эксперименты, подтверждают эффективность разработанных численных методов идентификации вероятностных моделей.

5. Структура диссертации

Диссертационная работа содержит 86 страниц машинописного текста, включает в себя введение, четыре главы, заключение и список литературы из 97 наименований.

6. Качество оформления

Оформление диссертации вполне соответствует установленным требованиям, имеются незначительные недочёты, которые не являются критическими.

7. Автореферат и опубликованные работы

Автореферат и публикации по теме диссертации в достаточной степени отражают содержание диссертационного исследования и научные положения, выносимые на защиту.

8. Замечания

По работе следует отметить ряд замечаний:

- На стр. 17 указано «обучение сети в ряде случаев приводит к тупиковым ситуациям; в таких случаях необходимо выбирать иные методы обучения сети, которые не всегда хорошо интерпретируемы». Неясно, что понимается под тупиковыми ситуациями;
- Из работы неясно, в каких единицах измеряется трудность тестового задания.
- На рис. 10 – вероятно перепутаны местами координатные оси, не указано, в каких единицах измеряется время. Также остаётся неясным, каким образом получены представленные распределения вероятностей.
- На рис. 12 несколько некорректно представлена схема структуры программного комплекса. С одной стороны ряд обозначений берётся из UML-диаграмм. Однако стрелки, скорее всего, обозначают не зависимости компонентов, а передачу данных. Этот факт нуждается в дополнительном пояснении.
- На стр. 31 упоминается, что модель включает в себя 8 состояний, однако в работе не описано, что обозначает каждое из этих состояний.
- На стр. 52 указано, что модель позволяет «оценивать пригодность тестовых заданий для измерения заявленных характеристик испытуемых». Как именно оценивается пригодность? Что понимается под «пригодностью».
- Из работе не совсем ясно, предусмотрен ли в ходе тестирования в режиме игровой диагностики переход в некоторое состояние, предполагающее, что дальнейшая работа системы невозможна, а испытуемый фактически не справился с этим и последующими заданиями. Возможно, есть какая-либо вероятность отказа испытуемого от прохождения теста?
- В разделе 3.4 - заявлены, возможности разработанных средств, но не приведено конкретных примеров.
- В работе отсутствуют выводы по отдельным главам.
- На стр. 69 упоминается «Рис. 17. Структура сформированного пула моделей для идентификации», однако, сам рисунок отсутствует.

9. Оценивая работу в целом

Считаю, что диссертация является законченной научно-квалификационной работой, в которой на основе проведённых автором исследований решена научная задача разработки подходов к оценке компетенций на основе процедур адаптивного компьютерного тестирования, учитывающих временную динамику выполнения тестовых заданий и изменения в состоянии испытуемых, а также особенности использования этих процедур в задачах игровой диагностики и оценки уровня подготовки операторов сложных технических систем.

Таким образом, рассматриваемая работа соответствует установленным

требованиям к кандидатским диссертациям, а её автор Думин Павел Николаевич заслуживает присуждения ему учёной степени кандидата физико-математических наук по специальностям 05.13.18 – Математическое моделирование, численные методы и комплексы программ, 05.13.01 – Системный анализ, управление и обработка информации (авиационная и ракетно-космическая техника).

Официальный оппонент

Аристов Антон Олегович,

119991, ГСП-1, Москва, Ленинский проспект., д.4

8-906-786-7124

e-mail : batan-87@mail.ru

доцент кафедры Автоматизации проектирования и дизайна

Национального исследовательского технологического университета «МИСиС»

13.11.2018

ПОДПИСЬ

Проректор по

научным во

д.т.у. «МИ

ЗАВЕРЯЮ

М. Исаев

